Example Guide
https://htmlpreview.github.io/?https://raw.githubusercontent.com/SDS-237-Data-Ethnography/sds-237-group-project-group-4/main/final-project.html?token=GHSAT0AAAAAABST4ULOFX54ODX7BDW5ZJSCZBBULVA

What is the Annual Survey of Jails Dataset?

The Annual Survey of Jails Dataset is collected and maintained by the Bureau of Justice Statistics (BJS) and Research Triangle Institute (RTI). The Purpose of the ASJ Dataset is to collect data on the characteristics of local jails and their inmates. The ASJ Dataset includes valuable information on the jail population including average daily population, rated capacity, supervised unconfined population, admissions, and releases, jail population by sex, race, and conviction status, juvenile population, and the number of inmates held for state and federal authorities.

Who made the dataset?

From 1982 to 2014, the Annual Survey of Jails Dataset was collected and cleaned by the Bureau of Justice Statistics (BJS). The BJS “builds and maintains respect and trust through clear public commitments to professional practice and transparency in all that they do [and] should actively engage with users in determining priorities for data collection and analysis, make their data available widely on an equal basis to all users, formally and informally, and fully inform users of the strengths and weaknesses of the data.” Members from this institution must be independent of political and other undue external influences that might affect the collection, analysis, or reporting processes of data and appear as manipulated for political or other purposes.

Since 2015, the ASJ Dataset has been collected by the Research Triangle Institute International (RTI). The mission of RTI is “an independent nonprofit research institute dedicated to improving the human condition”. Their vision is to address the world's most critical problems with science-based solutions in pursuit of a better future. RTI International was founded in 1958 by universities, government, and business leaders. In the work of ASJ, RTI staff have “developed working relationships with every eligible jail jurisdiction in the United States.” They aim to “generate national, state, and local data on correctional facility operations and inform strategies to improve the health, safety, and well-being of those who live and work in correctional facilities.” They have provided full transparency of how they collect the data including but not limited how they handle null values, outliers, how long it takes to complete the surveys of data collection, how they are improving the data collection, they are also transparent about their response rate.

BJS discontinued the data collection because of the inconsistency caused by the consensus that they conducted, so they gave this responsibility to RTI. BJS now supervises the data collection work done by RTI, to meet and be consistent with their standardization of data collection of ASJ in history.

What is the purpose of the dataset?

The purpose of data collection in jails is to account for demographic purposes to collect information on individuals' age, gender, ethnicity, and race. Data on the type of offenses that people are in jail for, often narratives are created with this data to say that the majority of a certain race group is in jail for one type of offense; this topic will be discussed further in later sections. Another purpose of this data set is to keep track of jail operations and examine the different aspects of jail conditions: jail capacity, staffing, and facility characteristics which can help the government plan for budgets and further planning. Demand for incarceration is calculated by the data on the average daily population of jails that ASJ helps collect. ASJ data can help provide trend data to observe changes in jail populations over time. Lastly, the ASJ data helps criminal justice researchers and the public understand the jail system and help with decisions on policy and planning to help reduce incarceration rates.

Data sources

BJS

The Bureau of Justice Statistics (BJS), under the U.S. Department of Justice, takes its primary role in collecting, analyzing, and disseminating statistical information on crime, criminal offenders, victims of crime, and the operation of justice systems at all levels of government. The Annual Survey of Jails (ASJ) is one of the programs conducted by the BJS.

The ASJ collects data on the demographic characteristics of jail inmates, their offenses, the case and court processing of these individuals, and other relevant information related to the jail system across states in the United States. The data collected through the ASJ provides insights into the characteristics and trends of jail populations for the government to take measures on the needs of inmates, staffing, faculty, etc.

The BJS contributes to the public's understanding of the criminal justice system in the US. Researchers, policymakers, and the public use this data to inform discussions, policies, and decisions related to the criminal justice system. The BJS serves as a key source for comprehensive statistics in this domain.

RTI

The Research Triangle Institute collaborates with government agencies, including the Bureau of Justice Statistics (BJS), to conduct surveys, gather data, and analyze information related to the jail system in the United States. This could include projects such as the Annual Survey of Jails (ASJ).

RTI also engages in evaluating the effectiveness of programs and interventions within the jail system. This could involve assessing the impact of various policies, rehabilitation programs, alternative sentencing initiatives, or ways to reduce incarceration rates. This is done through the research conducted by the RTI that informs the development of criminal justice policies; studying trends, identifying challenges, and proposing evidence-based solutions.

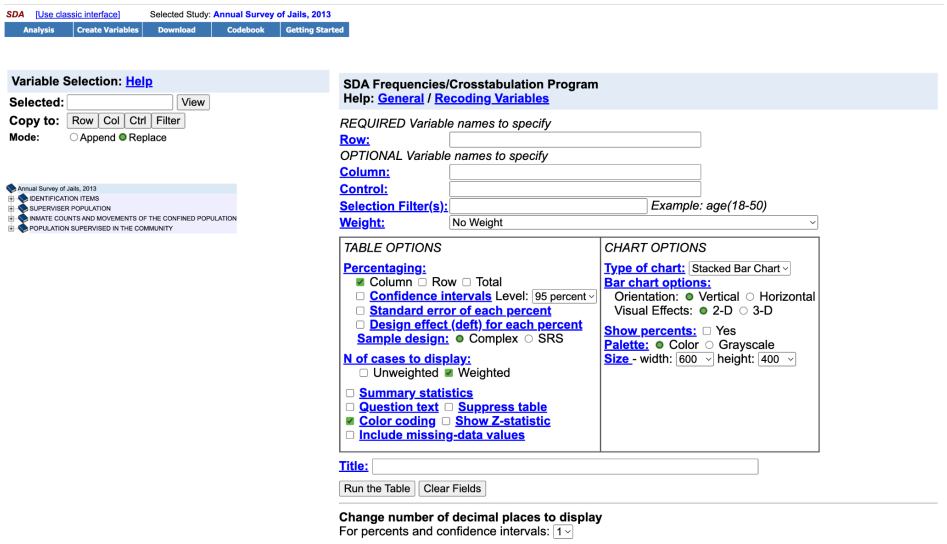
The RTI also provides technical assistance to government agencies, offering expertise in areas such as statistical analysis, research design, and evaluation methodologies.

How to use the ASJ dataset?

Users may access the Annual Survey of Jails (ASJ) dataset from the official website of the Bureau of Justice Statistics. For legality reasons, users will need to download locally the data themselves from ICPSR with an account. After the user logs in to the ICPSR, the user will see a whole list of datasets, each with a unique ID. For example, in the 2013 dataset, the ICPSR is 35517. Because the ASJ dataset has been collected since 1982 and the latest one was published in 2021, the user will need to select the specific year and navigate to the specific one. To make it easy, the user has the option to show or hide the summary for each year's dataset. By clicking the “more options” button, the user can navigate the datasets by adding different filters: time period, funding agency, restriction types, subject terms, data file format, recent releases, type of analysis, geography, and data availability. Besides filters, the user can also search by different categories related to datasets in general, such as Studies, Data-related publications, Variables, Series, and NACJD Websites, with specific filters applied to each category.

When the user clicks one specific dataset, the user will be directed to another webpage. On the webpage, the user would have access to data documentation, which is the specific explanation of the dataset, as well as project scopes, methodology, versions, and analysis information in detail. Moreover, the user could not only browse the dataset online with the button “Explore Data”, but also do the “SDA Analyze Online” feature(refer to the attached screenshot to see more details).

The last option is to directly download the dataset, and it would be helpful for the user to make sure to download the correct year edition as a delimited file (.tsv or tab-separated values).



For either option (browse online or download), it’s important to filter to the specific categories and the correct versions. The SDA Analyze online tool is a bit outdated, and the downloaded version is in the raw data format, so there’s no preference between one over the other and the user might choose whatever works best for them.

Data Structure?

Jails are local facilities under the jurisdiction of a city, local district, or country. Jails are short-term holding facilities. There is a slight ambiguity between jails and prisons. Jails are short-term holdings for those awaiting trial or sentencing and prisons are long-term holdings for those who are sentenced. The ASJ's definition of a jail is drawn at the jurisdictional level, so anything that is not at that level does not get reported.

Every jail is distinguished by Research Triangle Institute's (RTI) criteria for jails and represented with an ID unique to RTI (RTIID). So each row of jails are a unique observation in this dataset. Each observation is categorized by city, county, and region of jail in the country. The narratives we can draw from the data (variables) include but are not limited to the following: the ratio of race in the jails, the population of the jails, the number of jails in each city, etc. However, we can see that there still exists the singular voice of these narratives is the Bureau of Justice Statistics which conducted the Annual Survey of Jails. "Jail reporting units are identified by the 21-digit GID-government identification codes that the U.S. Census Bureau uses to identify governments in the United States based on the 2009 Census of Governments. The GID code identifies the general location of governments, and indicates whether they are state, county, city, place, or municipality organizations (p.5)."

Since the data has been collected for nearly 40 years until now, it's helpful to know its timeline of changes:
* In 1982, the data first began to be collected.
* In 1985, the ASJ was first published.
* In 1994, new sampling procedures were also introduced to minimize the standard errors of estimates.
* In 2015 and 2016, they reported the inmate count on the last weekend of June (midyear) and they did a detailed characteristics report of the total inmate count on December 31.
* In 2017, they reverted to mid-year references because there was a seasonal variation, as the jails typically held fewer inmates during the end of the year compared to mid-year.

There are some categorical variables that help structure the whole dataset. For example, JURISID allows the user to understand what Jurisdictional level the jail falls under; CITY divides by which city the jail is located in. The possible term is reflected by the reported cities in the US. COUNTY is categorized by which county the jail is located. The possible term is reflected by the reported counties in the US.

The main residual category is OTHEROFFSPEC. "On June 29, 2018, how many persons CONFINED in [your jail facilities / this facility], regardless of conviction status, had an offense type of... other?", and if prisoners have multiple charges, they're supposed to report their most serious offense. This means that if prisoners have multiple charges, especially if they're variable

in type, that variety in charges will not be reflected by the survey. This residual category might fail to accurately record the range of crime types of inmates.

Same as categorical variables, numeric variables like CONF_POP record the confined population at midyear in each prison. The following question was posed to each prison: “On June 29, 2018, how many persons under the supervision of your [jail jurisdiction / jail] were... CONFINED in [your jail facilities / this facility]?” However, the data documentation goes on to specify the people who should be included/excluded from this population. For instance, people who are working in community-based programs but return to the jail at night to sleep should be counted, but prisoners who are on long-term transfer to other jurisdictions are not considered part of the population.

The definitions for convicts are based on the goal of the survey designers to determine the population of prisons. Therefore, the survey makers were aiming to find the number of people who might reside in specific prisons. However, there was a social process through which “residence” was determined, part of which was the assertion that people who sleep at a place reside there, even if they don’t spend time there outside of sleeping. Defining “residents” was a social process.

Limitations of the dataset?

As mentioned above, the Bureau of Justice Statistics (BJS) stopped conducting the Annual Survey of Jails (ASJ) in 2015, instead overseeing the process but delegating the actual collection to the Research Triangle Institute (RTI). This was because of inconsistencies in the data collection processes. 2014 was not the first time problems with the ASJ have been identified.

One common critique of the ASJ is that the data collected is not representative or robust enough. Tracy Velázquez, a Senior Project Manager with Pew Charitable Trust, [critiqued that](https://thehill.com/opinion/criminal-justice/3845445-timely-jail-data-can-be-a-guiding-light-for-the-criminal-justice-system/) the data collected by BJS’s ASJ is outdated and spotty: “there’s limited information that they release on the people who are in jail via demographics. For example, there’s no information on admissions by race or ethnicity at all”. While it’s worth noting that Velázquez was, in that piece, promoting the Jail Data Initiative, an alternative source of statistics about jails, her critique is not uncommon. Many people with experience in the field, like Velázquez, find the BJS’s data was weak.

BJS has made efforts to address this critique, including changing the day of year when they collect their data from. However, this change was not enough (in fact, two years later, they reversed the change and went back to their original data collection methods. One of the biggest advantages touted over the ASJ by resources like the Jail Data Initiative is how much more information it collects (demographics of prisoners by admission to jails) and from how many more places (collecting data from almost a third of America’s jails).

Conclusion?

While there are other resources that collect data from more jails than the ASJ does, the fact is that there is a serious dearth of data about American prisons and their occupants. The existence of more datasets about jails, including and especially those which are as old as the ASJ, cannot be understated. Access to the dataset is intuitive and allows the user to filter to specific time periods, geographic locations, and other variables. Lastly, the collaboration between BJS and RTI has demonstrated that the ASJ is capable of adapting and evolving to improve on issues it might have.